FURTHER EVALUATION OF THE ACCURACY OF REINFORCER SURVEYS: A SYSTEMATIC REPLICATION

JOHN NORTHUP

LOUISIANA STATE UNIVERSITY

The present report evaluates the accuracy of a reinforcer survey by comparing the survey results to the results of subsequent reinforcer assessments for 20 children using a concurrent-operants arrangement to assess relative reinforcer preference. Total accuracy for the survey was determined to be approximately 57%. The results provide a systematic replication of Northup et al. (1996) with a much larger sample of children. A need for the development of more accurate and comprehensive reinforcer assessment methods for verbal children is discussed.

DESCRIPTORS: reinforcer assessment, functional assessment, surveys, attention deficit hyperactivity disorder

Self-reports of potential reinforcers continue to be widely used to develop treatments for the most common and pervasive childhood behavior problems. The use of reinforcer surveys to identify potential reinforcers has a long history in child behavior therapy. The practice appears to be well established, and continued use may be further supported by the ease and apparent efficiency of administration. Nevertheless, poor correspondence between verbal self-reports and subsequent behavior has been long noted and often demonstrated (e.g., Bernstein & Michael, 1990; Risley & Hart, 1968). Children's ability to accurately name events that will reinforce future behavior may be particularly questionable. Northup, George, Jones, Broussard, and Vollmer (1996) demonstrated that correspondence between the results of a reinforcer survey and a subsequent evaluation of reinforcement effects was no better than chance for 4 children with a diagnosis of attention deficit hyperactivity disorder (ADHD). However, the evaluation of only 4 subjects for a comparison of this type is a major limitation of the Northup et al. study.

The Northup et al. (1996) study has apparently not been replicated, and no other similar evaluations of the accuracy and treatment utility of reinforcer surveys could be found. The present report provides a systematic replication of the procedures of Northup et al. by (a) conducting a comparison of the results of a reinforcer survey and subsequent reinforcer assessments for a much larger sample of 20 children, and (b) presenting potential reinforcers (token coupons) simultaneously in a more efficient concurrent-operants arrangement rather than singly as in Northup et al. (1996). Presenting the token coupons simultaneously is both more efficient and provides a more direct assessment of relative reinforcement effects.

METHOD

Participants

The files of all children who attended a summer program for children with a diagnosis of ADHD during the past 5 years were reviewed. Children were included in this evaluation on the basis of the following criteria: (a) The results of a reinforcer survey as described below were available; (b) the results of a reinforcer assessment that included at least three baseline sessions, three assess-

ment sessions, and three return to baseline sessions were also available; and (c) the results had not been published elsewhere.

General Procedure

The survey and reinforcer assessment procedures were similar to those of Northup et al. (1996) with the following exceptions: (a) Token coupons were presented simultaneously in a concurrent-operants arrangement, (b) a separate category for peer attention was included, and (c) easy math problems, rather than a coding task, were used during the reinforcer assessment. Token coupons were used to represent a wide range of potential reinforcers that are common to this population but that are difficult or impossible to tangibly or immediately provide during an assessment (e.g., to get out of a school activity).

Reinforcer survey. The reinforcer survey included a total of 42 items, with seven items representing each of the following six categories; edible items, peer attention, activities, tangible items, teacher attention, and escape. The survey was administered verbally with the instruction that "I'm going to name some things that kids sometimes get in school. I want to know how much you like each of these things. After I name each thing you tell me if you like it 'not at all,' 'a little,' or 'a lot.' For example . . ." Rankings of each item were given the value of 1, 2, and 3, for *not at all, a little,* and *a lot,* respectively, and a percentage score was calculated for each category by dividing the summed score of the item rankings by the total possible score. Categories with a score of 75% or greater were considered to be high preference. The complete survey is available from the author.

Reinforcer assessment. During baseline, the child was seated across from an examiner, given a worksheet containing easy math problems (i.e., problems previously completed with greater than 90% accuracy), and given the instruction that "you can do as much

as you want, as little as you want, or none at all." The session ended if the child stated that he or she did not want to do any, said that he or she was done, or worked for 5 min. During the reinforcer assessment, token coupons representing each of the six reinforcer categories and the associated stimuli on the survey (and a control coupon) were placed directly above the worksheet. The stimuli associated with each coupon were reviewed and sampled by each child prior to the assessment. The child was given the same instructions as during baseline but was also told that if he or she completed a criterion number of problems he or she could stop and take a coupon of his or her choice and that he or she could earn as many coupons of any kind as he or she wanted. The criterion number of problems required to choose a coupon was determined individually based on the average number of problems completed during baseline and was marked on the worksheet prior to the session. Problems did not have to be completed correctly; however, the correct number of digits for each answer had to be provided (accuracy was typically 100%, in that problems were predetermined to be easy). Sessions ended as in baseline except that no time limit was imposed and the child was allowed to work (and choose coupons) until he or she stated that he or she was done. A return to baseline was conducted to demonstrate any reinforcement effects associated with the token coupons. The chosen coupons could be exchanged for any of the associated stimuli at any time upon request, as in Northup et al. (1996).

RESULTS AND DISCUSSION

The overall accuracy of the survey was determined as follows. For the survey, each category of stimuli (i.e., edible items, peer attention, etc.) was first determined to be either high (>75%) or low (<75%) prefer-

Table 1
Accuracy of the Reinforcer Survey

True positives	True negatives	False positives	False negatives	Total possibilities	Total accuracy
41 (34%)	28 (23%)	35 (29%)	16 (13%)	120	57%

ence for each participant. For the reinforcer assessment, the total number of problems associated with each coupon was first determined for each session (i.e., all problems completed prior to each coupon selection were attributed to the coupon selected). Subsequently, all token coupons associated with clear reinforcement effects were identified for each participant. The criteria for clear reinforcement effects were the same as in Northup et al. (1996) and were as follows: (a) The number of completed math problems associated with the coupon was greater than baseline, and (b) the number completed was greater than those for the control coupon for each session. The accuracy of the reinforcer survey then was calculated by determining the number of (a) true positives (categories identified as high preference on the survey that functioned as reinforcers), (b) false positives (categories that were identified as high preference that did not function as reinforcers), (c) true negatives (categories that were identified as low preference on the survey that did not function as reinforcers), and (d) false negatives (categories that were identified as low preference on the survey that did function as reinforcers). Total accuracy was calculated as the sum of true positives and true negatives divided by the total of all possibilities.

Table 1 shows the overall accuracy of the reinforcer survey compared to the results of the reinforcer assessments. The total accuracy of 57% was nearly identical to the 55% accuracy reported by Northup et al. (1996). The percentage of true and false positives and negatives was also quite similar and nev-

er varied by more than 7%. Overall, the results of this comparison replicate those of Northup et al. (1996) with a much larger sample of children. The results further suggest that reinforcer surveys may add little to an assessment beyond chance. The relatively high number of false positives and low number of false negatives again suggest that surveys may more accurately identify stimuli that are not reinforcers than those that are. False positive results would typically be expected to be of much greater concern than false negative ones in applied settings. That is, selecting stimuli for use in a behavioral treatment that do not in fact function as reinforcers (a false positive) will presumably result in treatment failure, whereas a failure to identify a potential reinforcer (false negative) may be of little or no concern (as long as at least one true reinforcer is identified). Because the results were very similar to those of Northup et al. (1996), they further suggest that presenting the token coupons simultaneously, rather than singly, had little or no effect on overall results.

A limitation of the procedures is that the potential effect of a single highly preferred item on the survey could have been masked by its inclusion in a category with other low-preference items (a false negative for only that item). However, an inspection of individual survey responses did not indicate this outcome, and a comparison of overall categories was of primary interest in this study. Similarly, it is possible that the absolute value of any one coupon may be obscured by the simultaneous presentation of all coupons. However, reinforcer assessments are

most likely to be conducted in applied settings when there is a particular interest in identifying the most potent or valued reinforcers for use in a behavioral treatment. It may also be considered a design limitation of this study that survey categories received a score independent of any comparison to any other category.

Some stimuli, such as access to widely popular edible items and toys, may typically function as reinforcers for so many children that no formal assessment is needed. However, the current challenge for behavior analysts is to accurately identify naturally occurring contingencies that maintain problem behaviors such as the relative value of teacher, parent, and peer attention, or escape from aversive events. What is needed is the further development of practical and efficient assessment methods for verbal children that can more accurately identify such contingencies. For example, the further development of verbal choice procedures and the evaluation of other procedures (e.g., instructions) or conditions that may mediate reinforcer assessment accuracy for verbal children may be fruitful directions for future research. The use of more experimentally based reinforcer assessments, such as those described in this report, may offer both a valid criterion measure for future research with verbal children and an interim alternative method of assessment if a more rigorous reinforcer assessment is considered necessary or desirable.

REFERENCES

Bernstein, D. J., & Michael, R. L. (1990). The utility of verbal and behavioral assessment of value. *Journal of the Experimental Analysis of Behavior, 54,* 173–184.

Northup, J., George, T., Jones, K., Broussard, & Vollmer, T. (1996). A comparison of reinforcer assessment methods: The utility of verbal and pictorial choice procedures. *Journal of Applied Behavior Analysis, 29,* 201–212.

Risley, T. R., & Hart, B. (1968). Developmental correspondence between the nonverbal and verbal behavior of preschool children. *Journal of Applied Behavior Analysis, 1,* 267–281.